Eigen-Inference Moments Method for CognitiveWireless Communications

Antonia Masucci, Oyvind Ryan, Sheng Yang, Merouane Debbah

Future Network and MobileSummit 2010 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2010
ISBN: 978-1-905824-16-8

Eigen-Inference Moments Method for Cognitive Wireless Communications

A.M. Masucci, Ø. Ryan, S. Yang, M. Debbah

Abstract: In many situations, telecommunication engineers are faced with the problem of extracting information from the network. This corresponds in many cases to infer on functionals of spectrum of random matrices with only a limited knowledge on the statistics of the matrix entries. Here, the inference on the spectrum of random matrices is realized by moments method. In its full generality, the problem requires some sophisticated tools related to free probability theory and the explicit spectrum (complete information) can hardly be obtained (except for some trivial cases). Results in the asymptotic case and in the finite case are presented and simulations show how the moments method approach can be applied in practice. Several still open problems in this field are also presented.

1. Introduction

Recent studies of the last decade have shown that future communication systems should be designed to adapt their parameters to the environment. This point of view, introduced by Joseph Mitola [1] , represents the heart of random cognitive networks, that can be thought of as self-organizing networks where terminals and base stations interact through cognitive sensing capabilities. The current development of microeletronics allows us to suppose that these wireless systems, for which the spectrum utilization will play a key role, will be realized in the near future. In doing so, it is of great importance to develop mathematical tools needed in this context. An interesting problem linked to random cognitive networks consists in understanding what can an intelligent device with n dimensions (time, frequency or space) extract in terms of useful information on the network from a set of K noisy measurements. Moreover, once this information has been extracted, how can the terminal exploit (by capacity estimation, power allocation, etc.) that information? In wireless cognitive networks, devices are autonomous and should take optimal decisions based on their sensing capabilities. The complexity of these systems requires some sophisticated tools such as free probability theory to make abstraction of the useless parameters. Free probability theory [2] was introduced by Voiculescu in the 1980s in order to attack some problems related to operator algebras and can be considered as a generalization of classical probability theory to noncommutative algebras. The analogy between the concept of freeness and the independence in classical probability leaves us to work with noncommutative operators like matrices that can be considered elements in what is called a noncommutative probability space. We consider of particular interest for our work, information measures such as capacity, signal to noise ratio and estimation of the signal power. Information measures are usually related to the spectrum (eigenvalues) of the channel matrix and not on the specific structure (eigenvectors). This entails many simplifications that make free probability theory, throught the concept of free deconvolution, a very appealing framework for the study of these networks. The general idea of deconvolution is related to the following problem [3]:

Given $\mathbf{A}$, $\mathbf{B}$ two $n \times n$ independent square complex Hermitian (or real symmetric) random matrices:

1) Can one derive the eigenvalue distribution of $\mathbf{A}$ from those of $\mathbf{A}+\mathbf{B}$ and $\mathbf{B}$? If feasible in the large n-limit, this operation is named additive free deconvolution,
2) Can one derive the eigenvalue distribution of $\mathbf{A}$ from those of $\mathbf{AB}$ and $\mathbf{B}$? If feasible in the large n-limit, this operation is named multiplicative free deconvolution.
The techniques generally used to compute the operation of deconvolution in the large n-limit are the moments method [3] and the Stieltjes transform method [4]. Each has its advantages and drawbacks. The moments method only works for measures with moments and characterizes the convolution only by giving its moments but it is easy implemantable and, in many applications, one needs only a subset of the moments depending on the number of parameters to be estimated. Instead, the Stieltjes transform method works for any measure and which allows, when computations are possible, to recover the densities. Unfortunately, this method works only in very few cases, since the operations which are necessary are almost always impossible to realize practically.
In this paper, we focus on the free deconvolution framework based on the moments method which uses the empirical moments of the eigenvalue distribution of random matrices to obtain information about the eigenvalues. The moments method has shown to be a fruitful technique in both the asymptotic and the finite setting to compute deconvolution. In the next section, the eigen-inference method will be analyzed in both the asymptotic and the finite setting. In the Section 3., we present an application showing how the moments method approach can be used. Several still open problems related to the moments method and some approaches to solve them are proposed in the Section 4.. In the following, upper (lower) boldface symbols will be used for matrices (column vectors), whereas lower symbols will represent scalar values. $(.)^H$ will represent the hermitian transpose operator. We let Tr be the trace for square matrices, defined by $\mathrm{Tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$, where a_{ii} are the diagonal elements of the $n \times n$ matrix $\mathbf{A}$. We also let tr be the normalized trace, defined by $\mathrm{tr}(\mathbf{A}) = \frac{1}{n}\mathrm{Tr}(\mathbf{A})$.

2. Eigen-Inference Moment Methods

The moments method [3] is based on the relations between the moments of the different matrices involved. For a given $n\times n$ random matrix $\mathbf{A}$, the p-th moment of $\mathbf{A}$ is defined, if it exists, as:

$$m_{\mathbf{A}}^{n,p} = \mathbb{E}\left[\mathrm{tr}(\mathbf{A}^p)\right] = \int \lambda^p d\rho_n(\lambda) \tag{1}$$

where $d\rho_n(\lambda) = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\delta(\lambda-\lambda_i)\right)$ is the associated empirical mean measure, and λ_i are the eigenvalues of $\mathbf{A}$.

2.1 Asymptotic Case

The use of the moments approach for the derivation of the eigenvalue distribution of random matrices dates back to the work of Wigner. Applied to single random matrix, the moments approach is shown to be a useful method for computing the eigenvalue distribution of classical known matrices, as semicircle law and Marchenko-Pastur law. When more than one matrix is considered, the concept of asymptotic freeness [2] leaves us to compute the eigenvalue distribution of sum and product of random matrices.

Definition 1. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be* $n \times n$ *hermitian random matrices and the functional* $\varphi(\mathbf{A}) := \lim_{n\to\infty} \frac{1}{n}\left[\mathrm{Tr}\left(\mathbf{A}\right)\right]$*, we said that* $\mathbf{A}$ *and* $\mathbf{B}$ *are aymptotically free if whenever there exist poly-*

nomials p_i and q_i such that $\varphi[p_i(\mathbf{A})] = 0$ for all i and $\varphi[q_j(\mathbf{B})] = 0$ for all j then

$$\varphi[p_1(\mathbf{A})q_1(\mathbf{B})p_2(\mathbf{A})q_2(\mathbf{B})\dots] = 0.$$

Given $\mathbf{A}, \mathbf{B}$ $n \times n$ hermitian and asymptotically free random matrices such that their eigenvalues distributions converge to some probability measure $\mu_{\mathbf{A}}$ and $\mu_{\mathbf{B}}$, respectivly, then the eigenvalue distributions of $\mathbf{A} + \mathbf{B}$ and $\mathbf{AB}$ converge to a probability measure which depends on $\mu_{\mathbf{A}}$ and $\mu_{\mathbf{B}}$, called additive and multiplicative free convolution, and denoted by $\mu_{\mathbf{A}} \boxplus \mu_{\mathbf{B}}$ and $\mu_{\mathbf{A}} \boxtimes \mu_{\mathbf{B}}$ respectively.
The idea of additive and multiplicative free convolution stems from the fact that:

$$m^p_{\mathbf{A}+\mathbf{B}} := \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\mathrm{Tr}\left((\mathbf{A}+\mathbf{B})^p\right)\right] = f(m^{(1)}_{\mathbf{A}}, \dots, m^{(p)}_{\mathbf{A}}, m^{(1)}_{\mathbf{B}}, \dots, m^{(p)}_{\mathbf{B}})$$

$$m^p_{\mathbf{AB}} := \lim_{n\to\infty} \frac{1}{n}\left[\mathrm{Tr}\left((\mathbf{AB})^p\right)\right] = f(m^{(1)}_{\mathbf{A}}, \dots, m^{(p)}_{\mathbf{A}}, m^{(1)}_{\mathbf{B}}, \dots, m^{(p)}_{\mathbf{B}})$$

which means that we can express the moments of $\mathbf{A} + \mathbf{B}$ and the moments of $\mathbf{AB}$ as a function of the moments of $\mathbf{A}$ and the moments of $\mathbf{B}$. In other words, the joint distribution of $\mathbf{A} + \mathbf{B}$ and the joint distribution of $\mathbf{AB}$ depend only on the marginal distributions of $\mathbf{A}$ and $\mathbf{B}$. Hence, when, for $n \to \infty$, the moment $m^{n,p}_{\mathbf{A}}$ converges almost surely to an analytical expression $m^p_{\mathbf{A}}$ that depends only on some specific parameters of $\mathbf{A}$ (such as the distribution of its entries)[1], one is able by recursion to express all the moments of $\mathbf{A}$ with respect only to the moments of $\mathbf{A}+\mathbf{B}$ and $\mathbf{B}$, or $\mathbf{AB}$ and $\mathbf{B}$. Since the distribution of $\mathbf{A} + \mathbf{B}$ ($\mathbf{AB}$) depends only on the probability measure associated with the moments of $\mathbf{A}$ and $\mathbf{B}$, one can define on the set of probability measures the operation of additive (multiplicative) free convolution.
Additive Free Deconvolution: The additive free deconvolution of a measure ρ by a measure ν is (when it exists) the only measure μ such that $\rho = \mu \boxplus \nu$. In this case, μ is denoted by $\mu = \rho \boxminus \nu$.
Multiplicative Free Deconvolution The multiplicative free deconvolution of a measure ρ by a measure ν is (when it exists) the only measure μ such that $\rho = \mu \boxtimes \nu$. In this case, μ is denoted by $\mu = \rho \boxslash \nu$.
We give the definition of non-crossing partitions.

Definition 2. *A partition π of $\{1, \dots, n\}$ is non-crossing if whenever we have four numbers $1 \le i < k < j < l \le n$ such that i and j are in the same block, k and l are in the same block, we also have that i, j, k, l belong to the same block. we denote by $NC(n)$ the set of non-crossing partition of $\{1, \dots, n\}$.*

The computation of free deconvolution in the asymptotic setting, by the moments method approach, is based on the moment-cumulant formula, which gives a relation between the moments $m^p_{\mathbf{A}} \equiv m^p_{\mu_{\mathbf{A}}}$ and the free cumulants $\kappa^p_{\mathbf{A}} \equiv \kappa^p_{\mu_{\mathbf{A}}}$ of a matrix $\mathbf{A}$, where $\mu_{\mathbf{A}}$ is the associated measure. It turns out that the cumulants are quantities much easier to compute, also thanks to the concept of non-crossing partitions . The moment-cumulant formula says that

$$m^p_{\mathbf{A}} = \sum_{\pi = \{V_1, \dots, V_k\} \in NC(p)} \prod_{i=1}^{k} \kappa^{|V_i|}_{\mathbf{A}}, \qquad (2)$$

[1]Note that in the following, when speaking of moments of matrices, we refer to the moments of the associated measure.

where $|V_i|$ is the cardinality of the block V_i. From (2) it follows that the first p cumulants can be computed from the first p moments, and vice versa. The following characterization enables to compute easily the additive free convolution using free cumulants.

Theorem 3. *[2] Given $\mathbf{A}$ and $\mathbf{B}$ asymptotically free random matrices, $\mu_{\mathbf{A}} \boxplus \mu_{\mathbf{B}}$ is the only law such that for all $p \geq 1$*

$$\kappa^p_{\mu_{\mathbf{A}} \boxplus \mu_{\mathbf{B}}} = \kappa^p_{\mu_{\mathbf{A}}} + \kappa^p_{\mu_{\mathbf{B}}} \tag{3}$$

Hence, the deconvolution $\mu_{(\mathbf{A}+\mathbf{B})} \boxminus \mu_{\mathbf{B}}$ of $\mu_{\mathbf{A}+\mathbf{B}}$ by $\mu_{\mathbf{B}}$ is characterized by the fact that for all $p \geq 1$

$$\kappa^p_{\mu_{\mathbf{A}+\mathbf{B}} \boxminus \mu_{\mathbf{B}}} = \kappa^p_{\mu_{\mathbf{A}+\mathbf{B}}} - \kappa^p_{\mu_{\mathbf{B}}}. \tag{4}$$

The implementation of additive free deconvolution is based on the following steps: for the two matrices $(\mathbf{A} + \mathbf{B})$ and $\mathbf{B}$, we first compute the free cumulants, then, considering the relation between the cumulants and the moments, we can obtain information about the distribution of the eigenvalues of $\mathbf{A}$.
The moments method, in the multiplicative case, is based on the relation between the moments $m^p_{\mathbf{A}} \equiv m^p_{\mu_{\mathbf{A}}}$ and the coefficients $s^p_{\mathbf{A}} \equiv s^p_{\mu_{\mathbf{A}}}$ of the S-transform of the measure associated to $\mathbf{A}$. They can be deduced one from each other from the following relations for all $p \geq 1$

$$m^1_{\mathbf{A}} s^1_{\mathbf{A}} = 1, \quad s^p_{\mathbf{A}} = \sum_{k=1}^{p+1} s^k_{\mathbf{A}} + \sum_{\substack{p_1, \ldots, p_k \geq 1 \\ p_1 + \cdots + p_k = p+1}} m^{p_1}_{\mathbf{A}} \ldots m^{p_k}_{\mathbf{A}}.$$

Hence, we can compute multiplicative free convolution by the following characterization.

Theorem 4. *[2] Given $\mathbf{A}$ and $\mathbf{B}$ asymptotically free random matrices, $\mu_{\mathbf{A}} \boxtimes \mu_{\mathbf{B}}$ is the only law such that:*

$$S_{\mu_{\mathbf{A}} \boxtimes \mu_{\mathbf{B}}} = S_{\mu_{\mathbf{A}}} S_{\mu_{\mathbf{B}}}$$

The multiplicative free deconvolution $\mu_{(\mathbf{AB})} \boxslash \mu_{\mathbf{B}}$ of $\mu_{\mathbf{AB}}$ by $\mu_{\mathbf{B}}$ is characterized by the fact that for all $p \geq 1$

$$s^p_{\mu_{\mathbf{AB}} \boxslash \mu_{\mathbf{B}}} s^1_{\mu_{\mathbf{B}}} = s^p_{\mu_{\mathbf{AB}}} - \sum_{k=1}^{p-1} s^k_{\mu_{\mathbf{AB}} \boxslash \mu_{\mathbf{B}}} s^{p+1-k}_{\mu_{\mathbf{B}}}. \tag{5}$$

In recent works, deconvolution, based on the moments method, has been analyzed when $n \to \infty$ for some particular matrices $\mathbf{A}$ and $\mathbf{B}$, such as when $\mathbf{A}$ is a random Vandermonde matrix and $\mathbf{B}$ is a deterministic diagonal matrix [6], or when $\mathbf{A}$ and $\mathbf{B}$ are two independent random Vandermonde matrices [7]. The authors in [6] developed analytical methods for finding moments of random Vandermonde matrices with entries on the unit circle and provide explicit expressions for the moments of the Gram matrix associated to the models considered. The explicit expressions found for the moments are useful for performing deconvolution. In these cases the moments technique has been shown to be very appealing and powerful in order to derive the exact asymptotic moments of "non free matrices".

2.2 Finite Case

In the finite setting (***i.e.***, when the sizes of the matrix are finite), the moments method intends to express in the additive case and in the multiplicative case

$$m_{\mathbf{A}+\mathbf{B}}^{n,p} = \frac{1}{n}\mathbb{E}\left[\mathrm{Tr}\left((\mathbf{A}+\mathbf{B})^p\right)\right] = f(m_{\mathbf{A}}^{(1)}, \ldots, m_{\mathbf{A}}^{(p)}, m_{\mathbf{B}}^{(1)}, \ldots, m_{\mathbf{B}}^{(p)})$$

$$m_{\mathbf{AB}}^{n,p} = \frac{1}{n}\mathbb{E}\left[\mathrm{Tr}\left((\mathbf{AB})^p\right)\right] = f(m_{\mathbf{A}}^{(1)}, \ldots, m_{\mathbf{A}}^{(p)}, m_{\mathbf{B}}^{(1)}, \ldots, m_{\mathbf{B}}^{(p)})$$

This means that we can express the moments of $\mathbf{A}+\mathbf{B}$ (or $\mathbf{AB}$) as a function of the moments of $\mathbf{A}$ and the moments of $\mathbf{B}$. Since the validity of the asymptotic assumption is rarely met in practice, towards a more realistic scenario, we study the moments method in the finite setting which provides tools to infer on some parameters of the channel matrix within a finite window of observation. Focusing on the study of random matrices in the finite case, the authors of [5] were able to derive the explicit series expansion of the eigenvalue distribution of various models, namely the case of non-central Wishart distributions as well as one sided correlated zero mean Wishart distributions. In particular, they proposed a general finite dimensional statistical inference framework based on the moments method in the finite case, which takes a set of moments as input and produces sets of moments as output with the dimensions of the matrices considered finite. They focus on the finite Gaussian case and, even though freeness usually does not hold for finite matrices, the moments method can still be used to propose algorithmic methods to compute these moments. The formulas of the moments presented in their contributions have been generated by iterations through partitions and permutations and use concepts from combinatorics. Similar results related to complex Wishart matrices are shown in [8], where exact formulas for moments and inverse moments of any order are obtained through the use of symmetric group.

The first and simplest result concerns the moments of a product of a deterministic matrix and a Wishart matrix. Let n, N be positive integers, $\mathbf{X}$ be $n \times N$ standard, complex, Gaussian[2] matrix and $\mathbf{D}$ a (deterministic) $n \times n$ matrix. Denoting the moments $D_p = \mathrm{tr}\left(\mathbf{D}^p\right)$ and $M_p = \mathbb{E}\left[\mathrm{tr}\left(\left(\mathbf{D}\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)^p\right)\right]$ for any positive integer p, Theorem 1 in [5] allows us to express the moments M_p in terms of the moments D_p. In particular, the first fourth moments can be written as

$$\begin{aligned} M_1 &= D_1 \\ M_2 &= D_2 + cD_1^2 \\ M_3 &= \left(1 + \frac{1}{N^2}\right) D_3 + 3cD_2D_1 + c^2D_1^3 \end{aligned}$$

where $c = \frac{n}{N}$. By a simple recursion, we can express D_p from M_p. For the first three moments these recursions become

$$\begin{aligned} &D_1 = M_1 \\ &D_2 = M_2 - cM_1^2 \\ &D_3 = \left(M_3 - 3c(M_2 - cM_1^2)M_1 + c^2M_1^3\right)\left(1 + \frac{1}{N^2}\right)^{-1}. \end{aligned}$$

[2] A standard complex Gaussian matrix $\mathbf{X}$ has i.i.d. complex Gaussian entries with zero mean and unit variance (in particular, the real and imaginary parts of the entries are independent, each with zero mean and variance $1/2$).

Considering the sum of a $\mathbf{D}$ deterministic $n \times N$ matrix and $\mathbf{X}$ a $n \times N$ standard, complex, Gaussian matrix, in accordance with the Theorem 2 in [5], for any positive integer p the moments $M_p = \mathbb{E}\left[\operatorname{tr}\left(\left(\frac{1}{N}(\mathbf{D}+\mathbf{X})(\mathbf{D}+\mathbf{X})^H\right)^p\right)\right]$ can be expressed in terms of the moments $D_p = \operatorname{tr}\left(\left(\frac{1}{N}\mathbf{D}\mathbf{D}^H\right)^p\right)$ as the following formulas:

$$\begin{aligned}
M_1 &= D_1 + 1 \\
M_2 &= D_2 + (2+2c)\,D_1 + (1+c) \\
M_3 &= D_3 + (3+3c)\,D_2 + 3cD_1^2 + \left(3 + 9c + 3c^2 + \frac{3}{N^2}\right) D_1 + \left(1 + 3c + c^2 + \frac{1}{N^2}\right)
\end{aligned}$$

where we have written only the first two moments. In this case also, by a simple recursion, one can express D_p from M_p. It is clear how the operation of deconvolution can be viewed as operating on the moments: explicit expression for the moments of the Gram matrices associated to our models (sum or product of a deterministic matrix and a complex standard Gaussian matrix) are found, and are expressed in terms of the moments of the matrices involved. Hence, deconvolution means to express the moments, in this case of the deterministic matrices, in function of the moments of the Gram matrices.
Similar results are found when the Gaussian matrices are assumed to be square and selfadjoint . The implementation of the results is also able to generate the moments of many types of combinations of independent Gaussian and Wishart random matrices.

3. Application

3.1 Power estimation

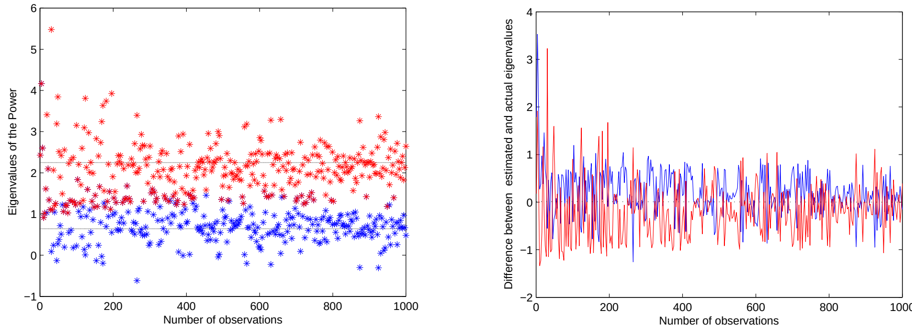
We consider a multi-user MIMO system where the received signal can be expressed by

$$\mathbf{y}_i = \mathbf{W}\mathbf{P}^{\frac{1}{2}}\mathbf{s}_i + \sigma \mathbf{n}_i \tag{6}$$

where $\mathbf{W}$, $\mathbf{P}$, $\mathbf{s}_i$, and $\mathbf{n}_i$ are respectively the $N \times K$ channel gain matrix, the $K \times K$ diagonal power matrix due to the different distances from which the users emit, the $K \times 1$ matrix of signals and the $N \times 1$ matrix representing the noise with variance σ. In particular, $\mathbf{W}, \mathbf{s}_i, \mathbf{n}_i$ are independent standard, complex, Gaussian matrices and vectors. We are interested in estimating the power, with which the users send information, from M observations (during which the channel gain matrix stays constant) of the vector $\mathbf{y}_i$. Considering the 2×2-matrix

$$\mathbf{P}^{\frac{1}{2}} = \begin{pmatrix} 1.5 & 0 \\ 0 & 0.8 \end{pmatrix} \tag{7}$$

and applying additive deconvolution first, and then multiplicative deconvolution twice (each application takes care of one Gaussian matrix), we can estimate the eigenvalues of $\mathbf{P}$ when we have an increasing number L of observations of the matrix $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_M]$, representing the signals received (we average across several block fading channels). Hence, we estimate the moments of the matrix $\mathbf{P}$ based only on the moments of the matrix $\mathbf{Y}\mathbf{Y}^H$. Knowing the moments of $\mathbf{P}$, we can estimate the eigenvalues using Newton-Girard formulas. When L increases, we get a prediction of the eigenvalues which is closer to the true eigenvalues of $\mathbf{P}$. Figure 1(a) illustrates the estimation of eigenvalues up to $L = 1000$ observations and Figure 1(b) shows that the difference between the estimated eigenvalues of the power and the actual ones tends to zero when the number L of observations increses. The actual powers are 2.25 and 0.64, the variance σ of the noise is assumed equal to 0.1.



(a) Estimation of the powers for the model (6), where the number L of observations is increased.

(b) Difference between the estimated eigenvalues of the power and the actual ones for the model (6).

4. Open Problems

From the previous section it is clear that the framework of deconvolution and, in particular, the moments play a key role in the study of cognitive wireless communication within a finite window of observation. In order to obtain more precise estimations, it is important to continue to develop increasingly sophisticated tools. In this Section we present still open problems related to the moments method approach. In Section 2.1, we have seen that under the assumption of asymptotic freeness the asymptotic moments of a product (or sum) of two random matrices can be expressed only with the asymptotic moments of each matrix. In the finite setting (Section2.2), explicit expressions of the moments can be found when the matrices considered are Gaussian, but we would like to be able to do the same when we consider matrices with a more involved structure than the Gaussian ones, such as Vandermonde and Toeplitz matrices. We are also interested in analyzing under which general setting and hypothesis can the joint eigenvalue distribution be expressed by its marginals, which means to find what are the minimum conditions to have separation of the moments. Another interesting problem consists in considering general function g of the matrices $\mathbf{A}$ and $\mathbf{B}$ and in finding as follows $\frac{1}{n}\mathbb{E}\left[\mathrm{Tr}\left(\mathrm{g}\left(\mathbf{A},\mathbf{B}\right)^p\right)\right] = h(m_{\mathbf{A}}^{(1)},\ldots,m_{\mathbf{A}}^{(p)},m_{\mathbf{B}}^{(1)},\ldots,m_{\mathbf{B}}^{(p)})$ how to express the moments of $\mathrm{g}\left(\mathbf{A},\mathbf{B}\right)$ in terms of the moments of $\mathbf{A}$ and the moments of $\mathbf{B}$. This enables us to consider more involved models than basic sum or product, and consequently more sophisticated wireless systems.

The formulas of the moments, in the finite setting, presented in [5] have been generated by traversing sets of partitions, however there may exist expressions for the same formulas which do not involve combinatorial concepts. The case presented in the work of G. Tucci [9], for instance, gives a closed form expression when the product of a positive definite matrix and a Wishart matrix is considered, *i.e.* $\mathbb{E}\left[\mathrm{Tr}\left(\left(\mathbf{X}^H\mathbf{A}\mathbf{X}\right)^p\right)\right]$ where $\mathbf{A}$ is a $n \times n$ positive definite matrix and $\mathbf{X}$ is a $n \times n$ complex standard Gaussian matrix. In [9] a more general result is given for every continous and measurable function $f : \mathbb{R}^+ \to \mathbb{R}$ such that $\int_0^\infty e^{-\alpha t}|f(t)|^2 dt < \infty$; explicit expression for the average like

$\mathbb{E}\left[\mathrm{Tr}\left(f\left(\mathbf{X}^H\mathbf{A}\mathbf{X}\right)\right)\right]$ are found.

We are interested in finding explicit expression for the averages of functionals of random matrices as $\mathbb{E}\left[\mathrm{Tr}\left(f\left((\mathbf{D}+\mathbf{X})(\mathbf{D}+\mathbf{X})^H\right)\right)\right]$, where $\mathbf{D}$ is a $n\times n$ determistic matrix, $\mathbf{X}$ is $n\times n$ complex standard Gaussian matrix, and the function f is continous bounded. As in [9], we want to attack the problem using tools from Representation Theory, such as the Schur polynomials, which are symmetric polynomials in the eigenvalues of the matrix argument. The difficulty linked to this approach is due to the fact that there is not yet an explicit expression for such polynomials computed in the sum of two matrices. Finding an explicit expression for this formula allows us to provide a new estimation of the ergodic capacity for the MIMO system with Ricean channel.

5. Conclusion

In this paper we have given a state of the art on the eigenvalue inference through the moments method approach. We have analyzed the use of moments method approach to compute the operation of deconvolution in the asymptotic setting and in the finite setting. Simulations are presented in order to show how the moments method can be applied in practice. Still open problems linked to the moments method in the finite setting are presented, and some approaches to solve them are proposed.

References

[1] J. Mitola III, Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio, Royal Institute of Technology (KTH) Stockholm, Sweden, May, 2000.

[2] F. Hiai and D. Petz, The semicircle law, free random variables and entropy, Mathematical Survey and Monograph, Providence, RI: American Mathematical Society, Vol. 77, 2000.

[3] F. Benaych-Georges and M. Debbah, Free Deconvolution: from Theory to Practice, submitted to IEEE Transactions on Information Theory, 2008.

[4] B. Dozier and J. Silverstein, On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices, J. Multivariate Anal., Volume 98, Number 4, pp. 678-694, 2007

[5] Ø. Ryan. A. Masucci, S. Yang, M. Debbah, Finite Dimensional Statistical Inference, submitted to Transaction on Information Theory, 2009.

[6] Ø. Ryan and M. Debbah, Asymptotic Behaviour of Random Vandermonde Matrices with Entries on the Unit Circle, IEEE Transactions on Information Theory, Volume 55, Number 7, pp. 3115-3148, 2009.

[7] Ø. Ryan and M. Debbah, Convolution arising from random Vandermonde matrices, submitted IEEE Transactions on Information Theory, 2009.

[8] P. Graczyk, G. Letac, H. Massam, The complex whishart distribution and the symmetric group, Ann. Statist. Volume 31, Number 1, pp. 287-309, 2003 .

[9] G. H. Tucci, A note on averages over random matrix ensembles, to be submitted http://arxiv.org/PS_ cache/arxiv/pdf/0910/0910.0575v3.pdf